[redacted]

E7.3-1 1.1 0.3
CR-135572

Interim Report

ORSER-SSEL Technical Report 15-73

COMPARISON OF PREPROCESSING AND CLASSIFICATION TECHNIQUES AS APPLIED TO MULTISPECTRAL SCANNER DATA

J. R. Hoosty and G. J. McMurtry

ERTS Investigation 082
Contract Number NAS 5-23133

INTERDISCIPLINARY APPLICATION AND INTERPRETATION OF ERTS DATA
WITHIN THE SUSQUEHANNA RIVER BASIN

Resource Inventory, Land Use, and Pollution

(E73-11103) COMPARISON OF PREPROCESSING AND CLASSIFICATION TECHNIQUES AS APPLIED TO MULTISPECTRAL SCANNER DATA Interim Report (Pennsylvania State Univ.) 4 p HC $3.00 CSCL 14E G3/13

N73-33260

Unclas
01103

Office for Remote Sensing of Earth Resources (ORSER)
Space Science and Engineering Laboratory (SSEL)
Room 219 Electrical Engineering West
The Pennsylvania State University
University Park, Pennsylvania 16802

Principal Investigators:

Dr. George J. McMurtry
Dr. Gary W. Petersen

Date: May 1973

COMPARISON OF PREPROCESSING AND CLASSIFICATION TECHNIQUES AS APPLIED TO MULTISPECTRAL SCANNER DATA

J. R. Hoosty and G. J. McMurtry

The objective of this research was to develop and compare various preprocessing and classification techniques for pattern recognition applications to multispectral scanner (MSS) data[1]. This work was performed during the pre-launch phase of the ERTS-1 study and, since no ERTS data was available, the data used was taken from a set of MSS data collected in 1969 by the University of Michigan aircraft over southeastern Pennsylvania. This body of data had been collected for the Federal Highway Administration and was made available to Penn State by Harold T. Rib.

Trainable classifiers implementing different discriminant functions were studied, and linear and quadratic discriminant functions were selected for implementation. Training was achieved by adjustment of parameters within the discriminant functions, based upon known sets of MSS observations (training sets). Eight different pattern classes (concrete, asphalt, Elk soil, Berks soil, grass, trees, crops, and Penn soil) were chosen for classification, with 50 to 60 patterns per class in the training set. Classifiers were categorized with respect to the type of training employed as well as discriminant function form. Two general types of training were conducted: parametric and nonparametric. Parametric classifiers train on the statistical parameters of the training set. The nonparametric classifiers assume a discriminant function with unknown coefficients. These coefficients are adjusted by a correction rule contingent upon the classification of the patterns in the training set. The four classifiers implemented were (see Appendix for program descriptions):

1. Parametric classifier with linear discriminant function (MINDIS).

---

[1]Hoosty, J. R. (1973) "A Preprocessing and Classification System for Remotely Sensed Multispectral Scanner Data," M.S. thesis, The Pennsylvania State University, University Park, Pa.

2. Parametric classifier with quadratic discriminant function (PARAM).

3. Nonparametric classifier with linear discriminant function (NPARMAP).

4. Nonparametric classifier with quadratic discriminant function (QUADMAP).

Principal components analysis and data normalization were chosen as the preprocessing methods to be implemented as options to the classification programs. The implemented classifiers were run using the following options: unpreprocessed data; principal components analysis using 13, 6, and 2 components; and normalized data. Comparisons were made between preprocessing and classifier results in the areas of separability of the training set, accuracy on the test set, computation speed, and overall appearance of the output site map.

All classifiers reached an acceptable level of separation as evidenced by training set classification, and accuracy as evidenced by test set classification. The technique of initially running classifiers using crude classes and then inspecting the site map in collaboration with aerial photographs and soil maps proved to be excellent for refining pattern classes. This procedure was a forerunner of the more formalized hybrid approach developed later. Considering computation time, the classifiers with the more complex discriminant functions (quadratic) were slower than those with less complex discriminant functions (linear). Nonparametric classifiers generally took longer during the training phase than did parametric classifiers. Assumptions of initial weight values near the pattern class means allowed the nonparametric classifier with linear discriminant function to train faster than the same classifier with the initial weight values assumed at a greater distance from the class means.

Principal components analysis provided a means of dimension reduction while maintaining an acceptable level of classification accuracy. Using the number of principal components (6) corresponding to 99 percent of the total variance, yielded class separation and classification accuracy comparable to using all the dimensions or principal components (13). However, in general, there was an obvious deterioration in class separation and classifier accuracy which accompanied dimension reduction below the value corresponding to 99 percent of the variance.

The results of data normalization as a preprocessing technique were not conclusive. However, in general the results indicated that classification with normalized data is comparable to, and in some instances better than, classification performed with unpreprocessed data. Some indication also exists that data normalization may remove unwanted noise.

Overall site map appearance was best for the classifiers employing linear discriminant functions; however, class boundaries were best defined by quadratic discriminant functions. Asphalt was the most misclassified class of the eight classes selected.

All classification training procedures used in this research were supervised methods; that is the user selected the different pattern classes himself before classifying the data. An unsupervised class selection technique could also be used as preprocessing for the basic classifiers described here. The dimension reduction and corresponding classification accuracy of principal components analysis should also be compared with other feature selection methods, such as divergence and Bhattacharya distance. Sequential pattern classification and feature selection may also be investigated for applications to remotely sensed data.

Judging on the basis of the factors of separability, accuracy, speed, site map appearance, and ease of implementation, the classifiers employing the linear discriminant function are comparable to -- and at times superior to -- classifiers using the quadratic discriminant function. There are certain instances where the quadratic discriminant function has a definite advantage, e.g., in defining class boundaries. A study of the statistical and physical nature of the data proved to be an excellent aid in the selection and implementation of classifiers and preprocessing techniques, and in the interpretation and analysis of corresponding results. It must be concluded that no one preprocessor/classifier combination is universally optimal. A knowledge of the physical aspects of the classification problem (ground truth) along with careful statistical analysis is essential for proper pattern recognition.

ORSER-SSEL Technical Report 15-73
The Pennsylvania State University
May 1973